



Virginia's New Standards of Learning Tests: At a Glance

The SOL tests measure student progress in achieving the new Standards of Learning in:

- ✦ English
- ✦ Mathematics
- ✦ Science
- ✦ History and Social Science (History, Geography, Civics, and Economics)
- ✦ Computer Technology

Tests given in grades 3, 5, 8 and in certain high school courses (Computer Technology SOL tests given only in grades 5 and 8)

Tests measure knowledge as well as critical thinking skills as required by the SOLs

Developed in collaboration with Virginia educators, the testing contractor, and the Department of Education

Field tested in Spring of 1997

First administered in Spring of 1998

Test administrations increased to twice a year beginning in 1998-99

Standards of Accreditation:

- ✦ Require SOL test performance be considered as one factor in promotion and retention in grades 3, 5, and 8
- ✦ Allow school divisions to consider SOL test performance in grading at the high school level
- ✦ Eventually require the passage of a certain number of SOL tests in order to receive a high school diploma



Standard of Learning Tests: What is Covered

Grade 3

- * **English** Grade K-3 SOLs
- * **Mathematics** Grade K-3 SOLs
- * **History and Social Science** Grade K-3 SOLs
- * **Science** Grade K-3 SOLs

Grade 5

- * **English** Grade 4-5 SOLs — reading, literature, and research, direct writing prompt/editing
- * **Mathematics** Grade 4-5 SOLs
- * **History and Social Science** Virginia studies since 1607
- * **Science** Grade 4-5 SOLs
- * **Technology** Grade 5 SOLs

Grade 8

- * **English** Grade 6-8 SOLs — reading, literature, and research, direct writing prompt/editing
- * **Mathematics** Grade 6-8 SOLs
- * **History and Social Science** Grade 5-7 SOLs
- * **Science** Grade 6 SOLs
Life Science
Physical Science
- * **Technology** Grade 8 SOLs

High School Course-Related Tests

- * **English** Grade 9-11 SOLs — reading, literature, and research, direct writing prompt/editing
- * **Mathematics** Algebra I
Algebra II
Geometry
- * **History and Social Science** World History to 1000 A.D. plus World Geography
World History from 1000 A.D. to Present plus World Geography
U. S. History
World Geography (under development)
- * **Science** Earth Science
Biology
Chemistry



Standards of Learning Tests: Major Implementation Activities

November 1996-March 1997

- * Notify school divisions of what SOLs are covered on which test
- * Conduct new regional meetings on the new program
- * Begin content review committee meetings
- * Develop initial test blueprints/specifications
- * Convene policy advisory committee
- * Develop manuals and procedures
- * Produce and distribute sample item brochure

April 28-May 2, 1997

- * Initial set of items field tested

Summer/Fall 1997

- * Review items for later field testing
- * Review data from first field test
- * Finalize blueprints for each test
- * Revise procedures for testing
- * Recommend final recommendations
- * Conduct additional bias review on certain high school course tests

Fall 1997

- * Limited field testing of items
- * Public test blueprints/sample items

Winter 1998

- * Test form development
- * Manual/procedures finalization

Spring 1998

- * First operational testing

- * First score reports — general information

Summer/Fall 1998

- * Committee develops recommended cut scores for “proficient” and “advanced” performance levels
- * Board of Education approves cut scores
- * Spring 1997 data re-scored using performance levels

1998-1999 School Year

- * Begin testing twice a year

2000-01 School Year

- * First 9th grade class for which the tests are a requirement

Spring 2004

- * First graduating class under new SOA requirements

On-going

- * New item development and field testing



Standards of Learning Tests: A Closer Look at Test Item Development

SOL test items are the product of intense scrutiny before they are used.

The Standards of Learning tests are anchored in the Standards of Learning. There are a total of 27 tests. SOL tests are primarily multiple-choice test items; however, English writing tests at grades 5, 8, and high school contain a direct writing component with multiple choice items.

Each test has been developed through a process which involves extensive review, field testing, and additional review of individual test items and item statistics. Harcourt Brace Educational Measurement, the test contractor, proposed test items to match the SOLs. A Content Review Committee was convened for each SOL test to review items for field testing, and to review and revise test blueprints, which guide test construction and use. A separate Bias Review Committee provided additional review of certain high school course test items for potential bias toward any student or groups of students.

Membership of the Content Review Committees includes classroom teachers, supervisors, and university professors.

The role of the Content Review Committee is to:

- ✱ review and revise test blueprints
- ✱ review items for field test before each administration
- ✱ review items and item statistics after each field test
- ✱ approve items for placement in the test item bank
- ✱ review and revise test forms before each use

Test items are analyzed by the Content Review Committees against four criteria:

SOL-Item match:

- ✱ Does this item measure the SOL it was designed to measure?

Appropriateness:

- ✱ Does this item appropriately measure content or skills that students in Virginia should be required to have by spring of the designated grade level or near the end of the course?

Difficulty:

- ✱ Is the difficulty of this item appropriate for the grade or course?

Bias:

- ✱ Is the item free from any content that would offend or unfairly penalize students on the basis of personal characteristics such as gender, ethnicity, religion, or socioeconomic status?



The first field tests of items were conducted on a statewide basis during April and May of 1997. The field test served to identify problems with test items before the items were used in an actual administration. Both the Content Review Committees and Bias Review Committee reviewed test items after field testing. After final acceptance of a test item by the committees, it was entered into the Virginia test item bank.

Harcourt Brace Educational Measurement placed test items on a proposed test form according to the specifications of the test blueprint. Each of the Content Review Committees approved the test form for use in testing. The SOL tests were administered for the first time in the Spring of 1998.

The Content Review Committees and Bias Committees continue to meet throughout the life of the testing program to apply this process to every item before it is placed on a test.

Standards of Learning Tests: A Closer Look at the Design of the Test

How are the Standards of Learning tests designed?

Each test is developed according to the blueprint established for the test. A test blueprint is a guide for test construction and use, and serves a number of purposes:

as a guide to test developers as they write test questions and construct SOL tests

as a guide to educators, parents and students in that they show:

- ✱ the SOLs covered by the test and which, if any, have been excluded
- ✱ which SOLs are assigned to each reporting category
- ✱ the number of test items in each reporting category and on the total test

- ✱ general information about how the test questions were constructed
- ✱ the materials that students are allowed to use while taking the test

Each time an SOL test is developed it will adhere to the requirements of the test blueprint. In this way teachers, students, and parents will know what is emphasized on the test. The test blueprints were sent to school divisions for distribution to teachers in December of 1997. They have been on the Department of Education's web site since January 1998.

How is the test blueprint organized?

Each blueprint contains the following information:

- 1. Test Development Guidelines:** guidelines used by Harcourt Brace Educational Measurement and the members of the Content Review Committees in developing the SOL tests. This section contains three parts:



- ★ General Considerations — lists general considerations that were used in developing the test as well as considerations specific to a particular content area.
- ★ Item Format — lists information on how items for the test are constructed.
- ★ Ancillary Materials — lists any materials (e.g., calculators, rulers, protractors, compasses, dictionaries) that students are allowed to use while taking each test.

2. Blueprint Summary Table: a summary of the blueprint which displays the following information:

- ★ reporting categories for each test
- ★ number of test items in each reporting category
- ★ Standards of Learning included in each reporting category. SOLs are identified by numbers and letters that correspond to the original SOLs document
- ★ SOLs which are excluded from the test
- ★ number of operational items on the test

- ★ number of field-test items on the test
- ★ total number of items (operational and field-test items) on the test

3. Expanded Blueprint: provides the same information as the Blueprint Summary Table except that the full text of each SOL is included. In addition, SOLs that are excluded from the test are categorized by the reason they were not included.

What is a reporting category?

Each test covers a number of SOLs. In the test blueprint, SOLs are grouped into categories that address related content or skills. These categories are labeled Reporting Categories. For example, a Reporting Category for the Grade 5 Mathematics test is “Computation and Estimation.” Each of the SOLs in this reporting category addresses computation using addition, subtraction, multiplication, or division, or require the student to estimate the answer to a problem. When the results of the SOL tests are reported, the scores will be presented in terms of scores for each Reporting Category and a total test score.

Are some SOLs assigned to more than one reporting category?

Learning concepts under a particular SOL are sometimes coded to different reporting categories. For example, the Science SOL 2.7a which deals with the effects that weather and seasonal changes have on the growth and behavior of living things is assigned to the reporting category “Life Processes and Living Systems” in the Grade 3 Science test. However, SOL 2.7b which deals with the effects of weather and seasonal changes on weathering and erosion of the land surface is assigned to the reporting category “Earth/Space Systems and Cycles”. Each concept is assigned to only one reporting category.

Why are some SOLs not tested on the SOL tests?

In some content areas, there are SOLs that do not lend themselves to multiple-choice testing. For example, in English, the oral language SOLs cannot be appropriately assessed in a multiple-choice format. In other cases, an SOL listed in one content area is covered by a similar SOL in another content area. For example, English SOL 4.9 which addresses the use of available technology to research a topic



is covered by a similar SOL at Grade 5 (Computer/Technology SOL C/T 5.3).

At the end of the blueprint for each test, the SOLs not tested are listed in “SOLs Excluded from Testing”. In the expanded blueprint the SOLs excluded from testing are categorized by the reason they are not being tested.

Will all SOLs listed in the blueprint be assessed each time the SOL tests are given?

Due to the large number of SOLs in a content area for a grade, *every* SOL will *not* be assessed on every SOL test form. By necessity, to keep the length of a test reasonable, each test will take samples from the SOLs within a reporting category. However, every SOL is eligible for inclusion on each form of an SOL test.

Standards of Learning Tests: Setting the Passing Scores

The Board of Education asked Virginia educators to make recommendations regarding passing scores.

The passing scores for the Standards of Learning tests are referred to as “cut scores.” They represent the number of test items a student has to answer correctly to be classified according to the achievement levels on the tests, listed below. A two-tiered committee structure has been put in place by the Board of Education to implement the process for setting the passing scores.

Eight Standard Setting Committees were given the responsibility to recommend a range of potential scores establishing three levels of student achievement:

- * Does not meet the standards (fail)
- * Proficient in the standards (pass)
- * Advanced attainment of the standards (pass)

In addition to the Standard Setting Committees, an oversight committee called the Standard Setting Advisory Committee for SOL Assessments has been appointed to ensure the integrity of the standard setting process.

The Eight Standard Setting Committees have been formed from nominations from school division superintendents, educational organizations, and the business community.

The members include teachers, curriculum specialists, and PTA representatives. Criteria for composition of the committees were:

- * geographic distribution
- * ethnic/racial balance
- * knowledge of the content area being assessed
- * teaching experience
- * experience with students with disabilities
- * experience with limited-English-proficient students



Standard Setting Committees were convened to make recommendations in the following content areas:

Content Area	Recommendations for SOL test passing score
Grade 3	Each of four tests in English, Science, Mathematics, and History and Social Science
Grade 5	English (reading, literature, and research only), Mathematics, Science, and History and Social Science
Computer Technology	Grade 5 and Grade 8 tests
Writing	Grades 5, 8, and high school
Mathematics	Grade 8 Mathematics, Algebra I, Algebra II, and Geometry
Science	Grade 8 Science, Earth Science, Biology, and Chemistry
History and Social Science	Grade 8 History, World History to 1000 A.D. plus World Geography, World History from 1000 A.D. to the Present plus World Geography, U.S. History
English	English: Reading, Literature, and Research in Grade 8, and high school course-related tests

The process for determining the passing scores for the Standards of Learning assessments is intended to be sound and structured.

The process which has been followed to develop the recommended ranges for cut scores is based on what students *should* know and uses the judgments of professional educators. It is structured on legally defensible procedures used by other states in high-stakes testing.

The role of the Standard Setting Advisory Committee for SOL Assessments is to advise the Board of Education on the appropriateness of the procedures used.

The Standard Setting Advisory Committee reviewed the appropriateness of the composition of each of the eight Standard Setting Committees and the degree to which standard setting procedures were followed as prescribed. The chair of each of the eight Standard Setting Committees is a school division superintendent. Each will serve on the oversight committee (the Standard Setting



Advisory Committee for SOL Assessments). The remaining members of the oversight committee have been selected from various constituents with a significant stake in the improvement of student academic achievement. These groups include the Chamber of Commerce; parents' and teachers' associations; school division superintendents; higher education representatives; elementary, middle, and high school principals; and the Virginia Education Association.

Standard of Learning Tests: A Closer Look at Setting Passing Scores

Standard Setting Committees used two accepted methods.

Ultimately the Board of Education must decide, for each of the 27 SOL tests, how many questions a student must get right in order to be deemed "Proficient" or "Advanced." The eight Standard Setting Committees have followed one of two systematic methods for developing their recommended ranges of scores. In early 1998 the Department of Education piloted two accepted methods for setting standards with two groups of teachers

within the state. The pilots resulted in the recommendation that the Modified-Angoff method be used to set the standards for all SOL tests except English: Writing. In the case of English: Writing, the recommendation was the Bookmark method.

The Standard Setting Advisory Committee approved these recommendations, recognizing they were generally accepted methods for standard setting which have been used in other states in similar testing programs. The recommended methods were also accepted as appropriate for the type of tests given.

The standard setting process in each of the eight Standard Setting Committees was led by a consultant from BETA, Inc. BETA, Inc. has many years of experience in the field of standard setting in various states where such tests are used in high-stakes situations for students.

The Modified-Angoff Method

The Modified-Angoff procedure has been widely used to set performance standards on multiple-choice tests for over 25 years. This was the method previously used by Virginia to establish the passing score points for the Literacy Passport Tests.

Generally, the process asks the committee members to discuss and describe the three levels of proficiency — "Does Not Meet," "Proficient," and "Advanced." Committee members then individually examine each test item being considered. (The SOL test items from the Spring 1998 administration of the test were considered.) Each member is asked to judge each item in terms of what proportion of 100 "barely proficient" students, or "barely advanced" students *should* get a particular item correct. Each member's judgments are converted to a proposed "Proficient" or "Advanced" cut point.

The entire committee then views the results of this first round of judgments. They see what the range of proposed passing scores would be and how many judges are at each of those proposed scores. Individual test items are then discussed and views and perspectives of committee members expressed. The diversity of the group allows the discussion to be focused on the many aspects of instruction in the state as well as the many types of students instructed in Virginia's schools.

At no time during this discussion is the actual performance of Virginia's students on these tests discussed. No actual scores from Spring 1998 are used in the standard setting process. Instead, the discussion focuses on what a student should be able to do rather than his or her current level of performance.



After this initial discussion, each committee member is asked to return to the actual test items again, and, in light of the discussion, make a second round of ratings. The results of this second round are discussed within the committee. After the second discussion, committee members are asked to return to the test items for a third time and make their final judgments. It is the product of these third round judgments that form the recommendations forwarded to the Board of Education for consideration.

The Bookmark Method

The Bookmark procedure was chosen to be used with the English: Writing tests. For the Standards of Learning English: Writing Tests, at grades 5, 8, and high school, students write a brief paper on a topic provided at the time of testing and answer multiple-choice questions. The number of multiple-choice questions answered correctly is added to the number of points earned on the paper to derive a total test score.

As with the Modified-Angoff method, the committee members discuss and describe the three achievement levels of "Does Not Meet," "Proficient," and "Advanced." The committee members then receive test booklets with the multiple-choice

test items ordered from easiest to hardest, interspersed with examples of student writing ordered from low writing performance to high writing performance. Each committee member is asked to independently place two "bookmarks" in the booklet of ordered items and papers. The first is placed where the questions and papers preceding it represent knowledge and skills which a student must have to be considered "barely proficient" in writing. The second "bookmark" is placed where the questions and papers preceding it represent the knowledge and skills a student must have to be considered "barely advanced."

The committee members are shown their individual placements as well as the range of all the committee members. As with the Modified Angoff, discussion ensues which focuses on explanations of their judgments and discussion of what placement of the "bookmarks" in various places means for student achievement. After this discussion, committee members are asked to return to their ordered items and papers and make their judgments again. After this second round is completed, the results of the group are again discussed. The process concludes with the third round where committee members are asked a final time to place their "bookmarks" in the ordered test items and papers to designate their individual cut points. It is the range that results from this third and final

round which forms the recommendation that will be forwarded to the Board of Education for consideration.

The Board of Education and Setting the Standard

The eight Standard Setting Committees are not recommending a specific set of cut scores to the Board of Education; rather, they are forwarding to the Board the range of scores at each performance level (proficient or advanced) that result from their third-round judgments. The Board of Education will then consider these recommended ranges when the members determine the two cut scores for each test. The Board's review process includes a series of meetings and hearings designed to place the committee recommendations before the public in order to receive broad comment.

The Board of Education has made a commitment to refrain from using the actual performance data of students on the Spring 1998 tests in making their decision. Rather, the Board will base the setting of student performance standards, not in terms of how many students will pass or fail, but on information and advice from Virginia educators about what students *should* know.